Randomised Algorithms: Supplementary Questions (draft!)

January 2023

1 Basic techniques

Exercise 1 [In expectation vs whp]

(a) Why would one want to use randomness in an algorithm?

(b) Explain the difference between guarantees that come "in expectation" and guarantees that come "with high probability".

(c) Compare quantitatively the outputs of the following python programs:

```python
import numpy as np

ans = []
for _ in range(20):
    m = 1
    while np.random.random() < 0.8:
        m += 1
    ans.append(m)
print(ans)
```

and

```python
import numpy as np

ans = []
for _ in range(20):
    n = 10000
    m = n
    loads = np.zeros(n)
    for _ in range(m):
        i1 = np.random.randint(0, n - 1)
        i2 = np.random.randint(0, n - 1)
        i_min = i1 if loads[i1] < loads[i2] else i2
        loads[i_min] += 1

    ans.append(np.max(loads) - m / n)
print(ans)
```

(*Answer*)

(a)

(b)

(c) The outputs from the first program are not very highly concentrated:

```
[3, 2, 6, 10, 16, 7, 1, 2, 10, 3, 21, 17, 1, 5, 18, 6, 6, 14, 4, 5]
```

The outputs from the second program are much more highly concentrated:

```
[2.0, 2.0, 2.0, 2.0, 2.0, 2.0, 3.0, 2.0, 2.0, 2.0, 2.0, 2.0, 2.0, 2.0, 3.0, 2.0, 2.0, 2.0, 2.0, 2.0]
```

Exercise 2 [For sufficiently large n] In several settings, we only care about what happens for sufficiently large n. Convince yourself that the following inequalities hold for sufficiently large n:

- $6n + n^2 \leq 2n^2$,
- $3n^2 - 6n \geq 2n^2$,
- $5n \log^2 n + 3n - 2n \log n \geq 4n \log^2 n$.

Exercise 3 [High probability events combine well] In this exercise, you will see that high probability events combine well. (This may make more sense after you have seen a few examples where we combine high probability events)

(a) Let A and B be two events with $\mathbf{Pr}\left[\, A \,\right] \geq 1 - n^{-c}$ and $\mathbf{Pr}\left[\, B \,\right] \geq 1 - n^{-c}$ for some constant $c > 0$. Then $\mathbf{Pr}\left[\, A \cap B \,\right] \geq 1 - 2n^{-c}$.

(b) How would you interpret in words the result in (a)?

(c) More generally for events $E_1, \ldots, E_m$ with $\mathbf{Pr}\left[\, E_i \,\right] \geq 1 - n^{-c}$, we have that

$$\mathbf{Pr}\left[\, \bigcap_{i=1}^{m} E_i \,\right] \geq 1 - n^{c} \cdot m.$$

(d) Let A and B be two events such that $\mathbf{Pr}\left[\, A \mid B \,\right] \geq 1 - n^{-c}$ and $\mathbf{Pr}\left[\, B \,\right] \geq 1 - n^{-c}$ for some constant $c > 0$. Then, $\mathbf{Pr}\left[\, A \,\right] \geq 1 - 2n^{-c}$.

(e) How would you interpret in words the result in (d)?

(f) Generalise the result in (d).

(*Answer*)

(a) We start by upper bounding the complement of this event using the union bound.

$$\mathbf{Pr}\left[\, \neg(A \cap B) \,\right] = \mathbf{Pr}\left[\, \neg A \cup \neg B \,\right] \leq \mathbf{Pr}\left[\, \neg A \,\right] + \mathbf{Pr}\left[\, \neg B \,\right] \leq n^{-c} + n^{-c} = 2n^{-c}.$$

Therefore, $\mathbf{Pr}\left[\, A \cup B \,\right] \geq 1 - n^{-c}$.

(b) "Two high probability events occur simultaneously with high probability probability."

(c) As in (a), we upper bound the probability of the complement,

$$\mathbf{Pr}\left[\, \neg \bigcap_{j=1}^{m} E_j \,\right] = \mathbf{Pr}\left[\, \bigcup_{j=1}^{m} \neg E_j \,\right] \leq \sum_{j=1}^{m} \mathbf{Pr}\left[\, \neg E_j \,\right] \leq m \cdot n^{-c}.$$

Therefore, $\mathbf{Pr}\left[\, \bigcap_{j=1}^{m} E_j \,\right] \geq 1 - m \cdot n^{-c}$. Note that for this to be high probability then $m \leq n^{-c'}$ for constant $c' < c$.

(d) "If A holds with high probability when high probability event B holds, then A holds with high probability."

(e) By the definition of conditional probability,

$$\mathbf{Pr}\left[\, A \,\right] = \mathbf{Pr}\left[\, A \mid B \,\right] \cdot \mathbf{Pr}\left[\, B \,\right] \geq (1 - n^{-c}) \cdot (1 - n^{-c}) = 1 - 2n^{-c} + n^{-2c} \geq 1 - 2n^{-c}.$$

(f) Consider the events $E_1, \ldots, E_m$, and further $\mathbf{Pr}\left[\, E_i \mid E_{i-1}, \ldots, E_1 \,\right] \geq 1 - n^{-c}$, for any $i \geq 1$ and $\mathbf{Pr}\left[\, E_1 \,\right] \geq 1 - n^{-c}$. Then,

$$\begin{aligned}\mathbf{Pr}\left[\, E_n \,\right] &= \mathbf{Pr}\left[\, E_m \mid E_{m-1}, \ldots, E_1 \,\right] \cdot \mathbf{Pr}\left[\, E_{m-1} \mid E_{m-2}, \ldots, E_1 \,\right] \cdot \ldots \cdot \mathbf{Pr}\left[\, E_2 \mid E_1 \,\right] \cdot \mathbf{Pr}\left[\, E_1 \,\right] \\ &\geq (1 - n^{-c})^m \geq 1 - m \cdot n^{-c},\end{aligned}$$

where the last inequality follows by Bernoulli's inequality (see Exercise 11).

Extended Note 1 [Indexed sets] Sometimes you will see $(E_i)_{i=1}^n$ instead of $E_1, \ldots, E_n$ or $(E_i)_{i=0}^\infty$ instead of $E_1, E_2, \ldots$. This is just for notational convenience.

Exercise 4 [Probability amplification] Consider an algorithm A for a minimisation problem. Further, assume that A succeeds in finding the minimum solution with probability at least $p > 0$, otherwise produces a solution with larger value.
We define the algorithm A' which repeats algorithm A, $t := \lceil \frac{c}{p} \cdot \log n \rceil$ times for some constant c (and n being a parameter of the given problem) and returns the minimum of these values. Then, A' has a success probability of at least $1 - n^{-c}$.

(*Answer*) The failure probability of A' is at most

$$(1-p)^t \leq e^{-pt} \leq e^{-p \cdot \lceil \frac{c}{p} \cdot \log n \rceil} \leq e^{-c \log n} = n^{-c}.$$

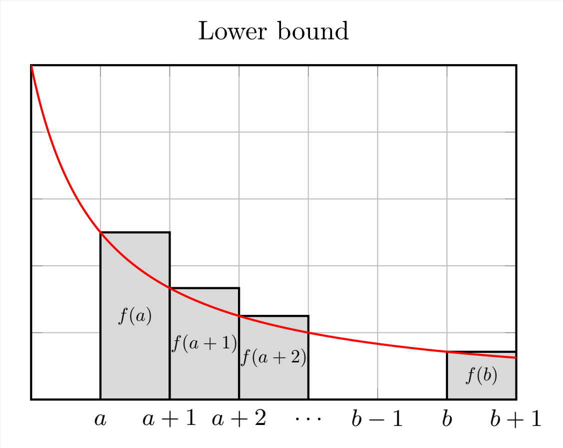
Hence, the success probability is at least $1 - n^{-c}$.

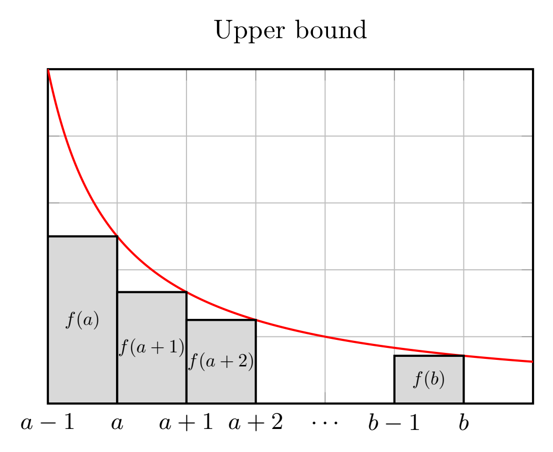
1.1 Common asymptotics

The following exercise demonstrates the general approach for bounding (discrete) sums using integrals. You may find it simpler to first attempt the bound on the harmonic numbers and then come back to this.

Exercise 5 [General approach] Let $f : [a-1, b+1] \to \mathbb{R}$ be a monotonically decreasing and integrable function for $a, b \in \mathbb{Z}$. Using the two figures below argue that

$$\int_a^{b+1} f(x)\,dx \leq \sum_{i=a}^{b} f(i) \leq \int_{a-1}^{b} f(x)\,dx.$$


Lower bound


Upper bound

(*Answer*) We begin with the observation that in both figures the sum $\sum_{i=a}^{b} f(i)$ is the total sum of the areas of the gray rectangles.
From the first figure, we see that the integral from a to $b+1$ is a lower bound for this area, since the function f is always below the rectangles

$$\int_a^{b+1} f(x)\,dx \leq \sum_{i=a}^{b} f(i).$$

Similarly, from the second figure, we see that f is always above the rectangles, so it is an upper bound for the sum

$$\sum_{i=a}^{b} f(i) \leq \int_{a-1}^{b} f(x)\,dx.$$

This concludes the proof.

Exercise 6 [Harmonic numbers] The n-th *harmonic number* H_n is defined as $H_n = \sum_{i=1}^n \frac{1}{i}$. Show that for any $n \geq 1$,

$$\ln n \leq H_n \leq \ln n + 1.$$

(*Answer*) We are going to apply the inequality from Exercise 5, for $f(x) = \ln(x)$ which is a decreasing function. The only problem is that for the upper bound we have that for $a = 1$, $1/(a-1)$ is undefined. We fix this by upper bounding the shifted sum as

$$\sum_{i=1}^{n} \frac{1}{i} = 1 + \sum_{i=2}^{n} \frac{1}{i} \leq 1 + \int_{1}^{n} \ln(x)\, dx = 1 + \ln(n) - \ln(1) = \ln(n) + 1.$$

For the lower bound we get,

$$\sum_{i=1}^{n} \frac{1}{i} \geq \int_{1}^{n+1} \frac{1}{x}\, dx = \ln(n+1) - \ln(1) = \ln(n+1) \geq \ln(n).$$

Further Reading 1 [Euler–Mascheroni constant] For most applications in algorithms a $\Theta(\log n)$ bound is sufficient for H_n (let alone a $\log n + \Theta(1)$ bound). In case you want to look it up,

$$\lim_{n\to\infty} H_n = \log n + \gamma = \log n + 0.577..,$$

where γ is known as the Euler–Mascheroni constant.

Exercise 7 [Factorial inequalities]

(a) Using Exercise 5, show that

$$n \log n - n + 1 \leq \ln(n!) \leq n \log n - n + 1 + \log n,$$

and deduce that

$$e \cdot \left(\frac{n}{e}\right)^n \leq n! \leq e \cdot n \cdot \left(\frac{n}{e}\right)^n.$$

(b) (optional) By drawing a figure and using the concavity of the $\log(\cdot)$ function, argue that

$$\int_{i-1}^{i} \log x\, dx \geq \frac{\log(i-1) + \log i}{2}.$$

By aggregating over all $i = 1, \ldots, n$ show that

$$n! \leq e\sqrt{n} \cdot \left(\frac{n}{e}\right)^n.$$

(*Answer*)

(a) The $\log x$ function is increasing in x, so we will apply the inequality of Exercise 5 to the function $f(x) = -\log x$, which is decreasing. Also, recall that

$$\int \log x\, dx = x \log x - x.$$

For the upper bound,

$$\sum_{i=2}^{n} -\log i \leq \int_{1}^{n} \log x\, dx = n \log n - n + 1.$$

For the lower bound,

$$\sum_{i=1}^{n} -\log i = -\log n + \sum_{i=1}^{n-1} -\log i \geq -\log n + \int_{1}^{n} \log x\, dx = -\log n + n \log n - n + 1$$

Combining the two bounds, we get that

$$n \log n - n + 1 \le \sum_{i=1}^{n} \log i \le n \log n - n + 1 - \log n.$$

By exponentiating both sides, we get that

$$e \cdot \left(\frac{n}{e}\right)^n \le n! \le e \cdot n \left(\frac{n}{e}\right)^n.$$

(b) Let $g(x) = \log x$, then $g'(x) = 1/x$ and $g''(x) = -1/x^2 < 0$, so the function is concave. Therefore, according to the following figure, the are of the trapezoid is contained in the area of the integral so we can deduce that

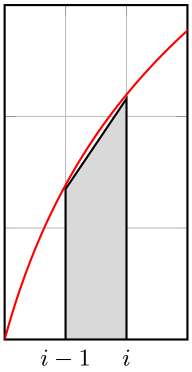
$$\int_{i-1}^{i} \log x \, dx \ge \frac{\log(i-1) + \log(i)}{2}.$$

By summing both sides for $i = 2, \ldots, n$, we get that

$$\int_{1}^{n} \log x \, dx \ge \sum_{i=1}^{n} \frac{\log(i-1) + \log(i)}{2} = \sum_{i=1}^{n} \log(i) - \frac{\ln(n)}{2}.$$

By exponentiating both sides, we get that

$$n! \ge e\sqrt{n} \cdot \left(\frac{n}{e}\right)^n.$$



Further Reading 2 [Stirling's approximation formula] Stirling's approximation formula says that

$$n! \sim n^n e^{-n} \sqrt{2\pi n},$$

and the following bounds hold for all n,

$$e^{1/(12n+1)} \le \frac{n!}{n^n e^{-n} \sqrt{2\pi}} \le e^{1/(12n)}.$$

1.2 Inequalities and Taylor estimates

In the course, you have often used the inequality $1 + x \le e^x$. In this section you wil prove that this inequality holds using a general (and simple) methodology.

Extended Note 2 Assuming that we want to prove the inequality

$$g(x) \le f(x),$$

for all values x in some range $[a, b]$.

- **Step 1:** Define the function
$$h(x) := f(x) - g(x).$$
- **Step 2:** Determine the minimum value of this function $h(x^*)$.
- **Step 3:** Deduce that
$$h(x) \geq h(x^*).$$
If $h(x^*) \geq 0$, then the inequality holds. Otherwise, for x^* the inequality fails.

Exercise 8 Prove the inequality $1 + x \leq e^x$ for any $x \in \mathbb{R}$.

(*Answer*) We proceed with the above steps:

- **Step 1:** Define the function
$$h(x) := e^x - (1 + x).$$
- **Step 2:** By differentiating we get
$$h'(x) = e^x - 1.$$
Setting $h'(x) = 0$, we get $x = 0$, which is a minimum since $h'(x) > 0$ for $x > 0$ and $h'(x) < 0$ for $x < 0$.
- **Step 3:** We can deduce that
$$h(x) \geq h(0) = 0,$$
or equivalently
$$e^x - (1 + x) \geq 0,$$
which gives us the claim
$$e^x \geq 1 + x.$$

Extended Note 3 [Inequalities via Taylor estimates] Another way to derive inequalities for sufficiently small x is via Taylor estimates. For instance,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots.$$

Then by truncating the series at a certain point, e.g.,

$$e^x \approx 1 + x + \frac{x^2}{2!},$$

we can obtain upper and lower bounds for sufficiently small x by changing the constant in the last term. For instance, we can obtain the inequalities,

$$e^x \leq 1 + x + x^2,$$

and

$$e^x \geq 1 + x + \frac{1}{4}x^2.$$

for any $|x| \leq 1/2$.

Exercise 9 [Birthday paradox] In this exercise, you will prove that with (positive) constant probability you need to take $\Theta(\sqrt{n})$ samples until a collision occurs. Let $\mathcal{E}_i$ be the event that the first collision occurs in the first i samples.

(a) Show that the probability that the first i samples are distinct is

$$\mathbf{Pr}\left[\, \mathcal{E}_i \,\right] = \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \ldots \cdot \left(1 - \frac{i-1}{n}\right).$$

(b) Using the inequality $1 + x \leq e^x$ and choosing a constant $C > 0$ appropriately show that

$$\mathbf{Pr}\left[\mathcal{E}_{C\sqrt{n}}\right] \leq 1/2.$$

(c) Using the inequality $e^x \leq 1 + x + x^2$ (for $x \leq 1$) show that $e^{-u-u^2} \leq 1 - u$ for $u \leq 1/2$.

(d) By choosing a constant $c > 0$ appropriately show that

$$\mathbf{Pr}\left[\mathcal{E}_{c\sqrt{n}}\right] \geq 1/e.$$

(e) State and deduce the birthday paradox. Why might it not (always) apply in the real-world?

(f) (Maybe) Looking at the crsids of the students that signed up for "Randomised algorithms", what can the birthday paradox say?

Exercise 10 [Lower bounding ratios] In this exercise, we will prove a simple inequality that is useful for bounding ratios, for instance in approximation algorithms.
For any $\epsilon \in (-1, 1)$, it holds that

$$\frac{1}{1+\epsilon} \geq 1 - \epsilon.$$

(a) Prove the inequality.

(b) Consider a maximisation problem P. Assume you have proven that the optimal solution is at most $n^2 + 10n$ and your algorithm returns a solution at least $n^2 - 3n \log n$. Then prove that your algorithm has an approximation ratio of at most $1 + o(1)$.

(c) Repeat the previous exercise when the lower bound is $\frac{1}{2}n^2 - 3n \log n$,

(*Answer*)

(a) Since $\epsilon \in (-1, 1)$, we have that $1 + \epsilon > 0$ and so the inequality is equivalent to

$$1 \geq (1 - \epsilon) \cdot (1 + \epsilon),$$

which is equivalent to

$$1 \geq 1 - \epsilon^2,$$

which holds by the assumption that $\epsilon \in (-1, 1)$.

(b) The approximation ratio is at most

$$r = \frac{n^2 + 10n}{n^2 - 3n \log n}.$$

By dividing by the dominant term on both numerator and denominator, we get that

$$r = \frac{1 + \frac{10n}{n^2}}{1 - \frac{3n \log n}{n^2}} = \frac{1 + \frac{10}{n}}{1 - \frac{3 \log n}{n}}.$$

Using the inequality from (a),

$$r \geq \left(1 + \frac{10}{n}\right) \cdot \left(1 + \frac{3 \log n}{n}\right) = 1 + \frac{10}{n} + \frac{3 \log n}{n} + \frac{30 \log n}{n^2} \leq 1 + \frac{4 \log n}{n}.$$

(c) Proceeding as in the previous question, we want to lower bound

$$r = \frac{n^2 + 10n}{\frac{1}{2}n^2 - 3n \log n}.$$

By dividing by the dominant term on the denominator, we get that

$$r = \frac{2 + \frac{20n}{n^2}}{1 - \frac{6n \log n}{n^2}} = \frac{2 + \frac{20}{n}}{1 - \frac{6 \log n}{n}}.$$

Using the inequality from (a),

$$r \geq \left(2 + \frac{20}{n}\right) \cdot \left(1 + \frac{6 \log n}{n}\right) = 2 + \frac{20}{n} + \frac{12 \log n}{n} + \frac{120 \log n}{n^2} \leq 1 + \frac{7 \log n}{n}.$$

Exercise 11 [Bernoulli's inequality] Prove that for any $n \in \mathbb{N}$ and for any $x \in [-1, +\infty)$, we have that

$$(1+x)^n \geq 1 + nx.$$

(*Answer*) We proceed by induction on n:

- **Base case:** For $n = 0$, we have that $1 \geq 1$, which is trivially true.
- **Inductive case:** Assume true for $n = k$, i.e.,

  $$(1+x)^k \geq 1 + kx.$$

  Then, for $n = k+1$ we have that

  $$\begin{aligned}(1+x)^{k+1} &= (1+x)^k \cdot (1+x) \\ &\geq (1+kx) \cdot (1+x) (\text{By I.H.}) \\ &= 1 + (k+1)x + kx^2 \\ &\geq 1 + (k+1)x (\text{since } x^2 \geq 0 \text{ and } k \geq 0).\end{aligned}$$

  Hence, it also holds for $n = k+1$.

1.3 Indicator random variables

Exercise 12 [Basic properties] Let A, B be two events in a sample space Ω. Then, prove the following properties for the indicators of these events:

(a) $\mathbf{E}[\mathbf{1}_A] = \mathbf{Pr}[A]$,

(b) $\mathbf{1}_{A \cap B} = \mathbf{1}_A \cdot \mathbf{1}_B$,

(c) $\mathbf{1}_{\neg A} = 1 - \mathbf{1}_A$,

(d) $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A \cap B}$,

(e) $\mathbf{Var}[\mathbf{1}_A] = \mathbf{Pr}[A] \cdot (1 - \mathbf{Pr}[A])$,

(f) $\mathbf{E}\left[(\mathbf{1}_A)^k\right] = \mathbf{Pr}[A]$ for any $k > 0$.

(*Answer*) Let $p = \mathbf{Pr}[A]$.

(a) By the definition of expectation,

$$\mathbf{E}[\mathbf{1}_A] = p \cdot 1 + (1-p) \cdot 0 = p.$$

(b) The indicator $\mathbf{1}_{A \cap B} = 1$ iff events A and B hold iff $\mathbf{1}_A = 1$ and $\mathbf{1} = B$ iff $\mathbf{1}_A \cdot \mathbf{1}_B = 1$.

(c) The indicator $\mathbf{1}_{\neg A} = 1$ iff $\neg A$ holds iff A does not hold iff $\mathbf{1}_A = 0$ iff $1 - \mathbf{1}_A - 1$

(d)

(e) Since $\mathbf{E}[\mathbf{1}_A] = p$ and since $\mathbf{Var}[X] = \mathbf{E}[(X - \mu)]$, we have that

$$\mathbf{Var}[\mathbf{1}_A] = p \cdot (1-p)^2 + (1-p) \cdot p^2 = p \cdot (1-p) \cdot (1 - p + p) = p \cdot (1-p).$$

(f) $\mathbf{1}_A^k = \mathbf{1}_A$ and so the conclusion follows from (a).

Exercise 13 [Inequalities]

(a) Argue that for random variables X and Y such that $X \leq Y$ (and have finite expectations), we also have that $\mathbf{E}[X] \leq \mathbf{E}[Y]$.

(b) (**Markov's inequality**) Let X be a non-negative random variable, then show that for any $a > 0$,

$$\mathbf{Pr}[X \geq a] \leq \frac{\mathbf{E}[X]}{a}.$$

Hint: Consider the indicator r.v. of the event $\{X \geq a\}$.

(c) (**Union bound**) Consider n events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ in a sample space Ω. Then,

$$\mathbf{Pr}\left[\bigcup_{i=1}^{n} \mathcal{E}_i\right] \leq \sum_{i=1}^{n} \mathbf{Pr}[\mathcal{E}_i].$$

Hint: See **Lecture 1 slide 12**.

(*Answer*)

(a) Let us define $Z := Y - X$. Then by the assumption $Z \geq 0$. Therefore,

$$\mathbf{E}[Z] = \int_{x=0}^{\infty} z \cdot p_Z(z)\, dz \geq 0,$$

since each term of the integrand is non-negative. Therefore,

$$\mathbf{E}[Z] \geq 0 \Rightarrow \mathbf{E}[Y - X] \geq 0 \Rightarrow \mathbf{E}[Y] - \mathbf{E}[X] \geq 0 \Rightarrow \mathbf{E}[Y] \geq \mathbf{E}[X].$$

(b) As the hint suggests, let $Z := \{X \geq a\}$. By a simple case distinction it follows that

$$X \geq Z \cdot a,$$

since

- **Case 1 [**$Z = 0$**]:** $X \geq 0 = Z \cdot a$ which trivially holds.
- **Case 2 [**$Z = 1$**]:** $X \geq a = Z \cdot a$, by the definition of the event Z.

Hence, by (a)

$$\mathbf{E}[X] \geq \mathbf{E}[Z \cdot a] = \mathbf{E}[Z] \cdot a = \mathbf{Pr}[X \geq a] \cdot a.$$

By rearranging, we get Markov's inequality.

(c) Let Z_i be the indicator for event $\mathcal{E}_i$ and let Z be the indicator for $\cup_{i=1}^{n} \mathcal{E}_i$. Then we have that

$$Z \leq \sum_{i=1}^{n} Z_i,$$

since $Z = 1$ iff any of the Z_i's is 1. Hence, by (a) we have that

$$\mathbf{E}[Z] \leq \sum_{i=1}^{n} \mathbf{E}[Z_i] \Rightarrow \mathbf{Pr}\left[\bigcup_{i=1}^{n} \mathcal{E}_i\right] \leq \sum_{i=1}^{n} \mathbf{Pr}[\mathcal{E}_i].$$

Exercise 14 [Fixed points of random permutation] The n passengers of an airplane take a seat uniformly at random.

(a) What is the number of passengers X that sit in their assigned seat in expectation?

(b) What is the variance of X?

(c) (optional +) Assume that n is divisible by 6, the passengers have arrived in $n/2$ pairs and there are $n/6$ rows of seats each with 6 seats. What is the expected number of pairs that will seat next to each other?

(*Answer*)

(a) Let X_i be the indicator of the event $\mathcal{E}_i$ that the i-th person sat in the correct seat. Then,

$$\mathbf{Pr}\left[\, \mathcal{E}_i \,\right] = \frac{1}{n}.$$

Therefore, for the total number of people that sat in their correct seat $X = \sum_{i=1}^n X_i$, by linearity of expectation we have that

$$\mathbf{E}\left[\, X \,\right] = \sum_{i=1}^n \mathbf{E}\left[\, X_i \,\right] = \sum_{i=1}^n \mathbf{Pr}\left[\, \mathcal{E}_i \,\right] = n \cdot \frac{1}{n} = 1.$$

Therefore in expectation one person will take the correct seat.

(b) Using the formula for the variance

$$\begin{aligned} \mathbf{Var}\left[\, X \,\right] &= \mathbf{E}\left[\, X^2 \,\right] - (\mathbf{E}\left[\, X \,\right])^2 \\ &= \mathbf{E}\left[\left(\sum i = 1^n X_i\right)^2\right] - 1 \\ &= \sum_{i=1}^n \sum_{j=1}^n \mathbf{E}\left[\, X_i X_j \,\right]. \end{aligned}$$

We distinguish two cases for $\mathbf{E}\left[\, X_i X_j \,\right]$:

- **Case [$i = j$]:** By the properties of indicators

$$\mathbf{E}\left[\, X_i^2 \,\right] = \mathbf{E}\left[\, X_i \,\right] = \frac{1}{n}.$$

- **Case [$i \neq j$]:** This one is a bit more involved:

$$\begin{aligned} \mathbf{E}\left[\, X_i X_j \,\right] = \mathbf{Pr}\left[\, X_i X_j = 1 \,\right] &= \mathbf{Pr}\left[\, X_i X_j = 1 \mid X_i = 1 \,\right] \cdot \mathbf{Pr}\left[\, X_i = 1 \,\right] \\ &= \mathbf{Pr}\left[\, X_j = 1 \mid X_i = 1 \,\right] \cdot \mathbf{Pr}\left[\, X_i = 1 \,\right] \\ &= \frac{1}{n-1} \cdot \frac{1}{n}, \end{aligned}$$

since given that index i is a fixed point, there are $n - 1$ remaining items in $n - 1$ slots, so the probability that j is a fixed point is $1/(n - 1)$.

By combining the two cases, we have that

$$\mathbf{Var}\left[\, X \,\right] = n \cdot \frac{1}{n} + n \cdot (n-1) \cdot \frac{1}{n \cdot (n-1)} - 1 = 1.$$

Exercise 15 [Inversions in a random permutation] Consider a permutation π of the set $[n] := \{1, \ldots, n\}$. An *inversion* is a pair (i, j) such that $i < j$ and $\pi_i > \pi_j$. Intuitively, it captures the pairs of indices that are out of order.

Show that if π is chosen uniformly at random, then the expected number of inversions N, satisfies

$$\mathbf{E}\left[\, N \,\right] = \frac{n \cdot (n-1)}{4}.$$

How does this quantity relate to *insertion sort*?

(*Answer*)

(a) We define the indicator random variable $X_{ij} = \{\pi_i > \pi_j\}$ (for $i < j$). By considering the $n!$ possible permutations of $[n]$, in half of them $\pi_i > \pi_j$ and in the other half $\pi_i < \pi_j$. So,

$$\mathbf{Pr}\left[\, X_{ij} \,\right] = \frac{1}{2}.$$

Then, the number of inversions $X = \sum_{i=1}^n \sum_{j=i+1}^n X_{ij}$ satisfies

$$\mathbf{E}\left[\, X \,\right] = \sum_{i=1}^n \sum_{j=i+1}^n \frac{1}{2} = \binom{n}{2} \cdot \frac{1}{2} = \frac{1}{4} \cdot n(n-1).$$

(b) This is the expected number of swaps performed by the insertion sort algorithm on a randomly permuted input.

Question: What should we do if the input is not random? We can randomly permute the input vector in $\mathcal{O}(n)$ time.

Exercise 16 [Records of a random permutation] Given a random permutation π of the set $[n] := \{1, \ldots, n\}$, how many times does line 5 in the following code execute?

```
1: function FindMin($x_1, \ldots, x_n$)
2:     $m \leftarrow \infty$
3:     for $i = 1$ to $n$ do
4:         if $x_i < m$ then
5:             $m \leftarrow x_i$
6:         end if
7:     end for
8:     return $m$
9: end function
```

(*Answer*) Let X_i be the indicator of the event $\mathcal{E}_i$ that the i-th element is larger than all previous elements $1, \ldots, i-1$. There are $i!$ possible permutations for elements $x_1, \ldots, x_i$ and in exactly $1/i$ of them (due to symmetry) the smallest element will be x_i. Hence,

$$\mathbf{Pr}\left[\, X_i \,\right] = \frac{1}{i}.$$

Therefore, we can deduce that the total number of times $X = \sum_{i=1}^{m} X_i$ that we execute line 5 is in expectation

$$\mathbf{E}\left[\, X \,\right] = \sum_{i=1}^{n} \mathbf{E}\left[\, X_i \,\right] = \sum_{i=1}^{n} \frac{1}{i} = H_n,$$

where H_n is the n-th Harmonic number (and so we also know that $\mathbf{E}\left[\, X \,\right] = \log n + \Theta(1)$.

Exercise 17 [Local max in random permutation] In a sequence $x_1, \ldots, x_n$ a local maximum is an index i such that x_i is at least as large as its (at most two) neighbours x_{i-1} and x_{i+1}.
Compute the expected number of local maxima in expectation for a random permutation of the elements $[n] := \{1, \ldots, n\}$.

(*Answer*) Let Z_i be the indicator for whether the i-th index is a local maximum. Then for i not being a boundary value (i.e., $1 < i < n$) we have that

$$\mathbf{E}\left[\, Z_i \,\right] = \mathbf{Pr}\left[\, X_{i-1} \leq X_i \leq X_{i+1} \,\right] = \frac{1}{3},$$

since out of the 3! random permutations there are exactly two which have i as a local maximum:

1	2	3
1	**3**	**2**
2	1	3
2	**3**	**1**
3	1	2
3	2	1

For the two boundary values each has a probability 1/2 of being a local maximum, so in total we have

$$\mathbf{E}\left[\, X \,\right] = \frac{n-2}{3} + \frac{1}{2} + \frac{1}{2} = \frac{n+1}{3}.$$

Exercise 18 [Number of cycles in random permutation] In this exercise, you will bound the number of cycles in a random permutation of the elements of the set $[n] := \{1, \ldots, n\}$.

(a) Consider a set of distinct elements $\{i_1, \ldots, i_k\}$ of $[n]$. Prove that the probability they form a cycle is

$$\frac{1}{n} \cdot \frac{1}{n-1} \cdot \ldots \cdot \frac{1}{n-k+1}.$$

(b) Prove that the expected number of cycles of length k in a random permutation is $1/k$.

(c) Deduce that the total number of cycles is $\ln n + \Theta(1)$.

(*Answer*)

(a) Let $Z_{i_1,\ldots,i_k}$ be the indicator of the event that $i_1, \ldots, i_k$ form a cycle. Then, each outgoing edge of i_j should point to i_{j+1} (or i_1 if $j = k$) and so

$$\mathbf{Pr}\left[\, Z_{i_1,\ldots,i_k} = 1 \,\right] = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \ldots \cdot \frac{1}{n-k+1}.$$

(b) Let X_k be the number of cycles of length k in a random permutation. There are $\frac{n!}{(n-k)!}$ ways of ordering k elements from $[n]$ and therefore $\frac{1}{k} \cdot \frac{n!}{(n-k)!}$ ways of obtaining cycles of k elements from $[n]$. Therefore,

$$\mathbf{E}\left[\, X_k \,\right] = \frac{1}{k} \cdot \frac{n!}{(n-k)!} \cdot \frac{1}{n} \cdot \frac{1}{n-1} \cdot \ldots \cdot \frac{1}{n-k+1} = \frac{1}{k}.$$

(c) By summing over all lengths from $k = 1, \ldots, n$, we get that the expected number of cycles is

$$\mathbf{E}\left[\, X \,\right] = \sum_{k=1}^{n} \mathbf{E}\left[\, X_k \,\right] = \sum_{k=1}^{n} \frac{1}{k} = H_n = \log n + \Theta(1).$$

Exercise 19 [Monte-Carlo Estimation] In *Monte-Carlo* estimation algorithms, the goal is to estimate the volume of a body by sampling. For instance, to estimate the value of π, we can randomly sample points in the $[-1, 1]^2$ square and count which ones are within the circle of radius 1 centred at the origin. Then we report the area to the be the count over four times the number of samples taken.

(a) Show that this algorithm produces the correct answer in expectation.

(b) What kind of guarantees can you get by applying Markov's inequality?

(c) What kind of guarantees can you get by applying Chebyshev's inequality?

(d) What kind of guarantees can you get by applying the Chernoff bound?

(e) How would you modify the algorithm so that you only sample from the square $[0, 1]^2$?

(*Answer*) Let X_i be the indicator of the event $\mathcal{E}_i$ that the i-th sample is within the circle of radius 1. The reported value is $X := \frac{4}{n} \sum_{i=1}^{n} X_i$.

(a) The area of the square is 4 and the are of the circle is $\pi/4$. Hence,

$$\mathbf{Pr}\left[\, \mathcal{E}_i \,\right] = \frac{\pi}{4}.$$

Hence, by taking the average of n samples we have that

$$\mathbf{E}\left[\, X \,\right] = \frac{1}{n} \sum_{i=1}^{n} \mathbf{E}\left[\, X_i \,\right] = \frac{4}{n} \cdot n \cdot \frac{\pi}{4} = \pi.$$

(b) By Markov's inequality, we get that

$$\mathbf{Pr}\left[\, X \geq 3\pi \,\right] \leq \frac{1}{3}.$$

We can also obtain a lower bound by defining $Y := 4 - X$. Then

$$\mathbf{Pr}\left[Y \geq \frac{3}{2} \cdot \mathbf{E}\left[\, Y \,\right] \right] \leq \frac{2}{3},$$

and therefore,

$$\mathbf{Pr}\left[Y < \frac{3}{2} \cdot (4-\pi) \right] \geq \frac{1}{3}.$$

By replacing X we have that

$$\mathbf{Pr}\left[\frac{3}{2}\pi - 2 < X \right] \geq \frac{1}{3}.$$

So with probability at least 1/3, $\hat{\pi} \in [2.712, 9.41]$.

(c) By Exercise 12 (e) and since the X_i's are independent, we have that

$$\mathbf{Var}\left[X \right] = \frac{4^2}{n^2} \sum_{i=1}^{n} \mathbf{Var}\left[X_i \right] = \frac{4}{n} \cdot \pi \cdot \left(1 - \frac{\pi}{4} \right) = \frac{\pi(4-\pi)}{n}.$$

Therefore, by Chebyshev's inequality, we have that

$$\mathbf{Pr}\left[|X - \pi| \geq k \cdot \sqrt{\frac{\pi(4-\pi)}{n}} \right] \geq \frac{1}{k}.$$

(d) Let $\tilde{X} = \sum_{i=1}^{n} X_i$. Then, by the Chernoff bound, we have that

$$\mathbf{Pr}\left[|\tilde{X} - \frac{\pi n}{4}| \geq t \right] \leq 2 \cdot e^{-2t^2/n}.$$

By scaling by $4/n$ we get

$$\mathbf{Pr}\left[|X - \pi| \geq \frac{4t}{n} \right] \leq 2 \cdot e^{-2t^2/n}.$$

Choosing $t := k\sqrt{n}$, we get that

$$\mathbf{Pr}\left[|X - \pi| \geq \frac{4t}{n} \right] \leq 2 \cdot e^{-2k^2},$$

which is much better than the dependence in Chebyshev's inequality.

Exercise 20 [3-CNF] We are given a CNF formula consisting of m clauses, i.e., an expression of the form $C_1 \wedge C_2 \wedge \ldots \wedge C_m$ with $C_i = X_{i_1} \vee \ldots \vee X_{i_{c_i}}$. Consider a simple algorithm that randomly assigns a value to each of the literals X_i.

(a) Show that when $c_i = 3$ for all $i \in [m]$, then the output contains $\frac{7m}{8}$ satisfied clauses in expectation.

(b) Find a similar expression for the general case.

(*Answer*)

(a) Let X_i be the indicator variable for the event $\mathcal{E}_i$ that the i-th clause is satisfied. Then the clause will be satisfied if either of the three literals is true. This happens with probability $1 - 1/8$. Hence, for $X = \sum_{i=1}^{n} X_i$ being the total number of satisfied clauses, we have that

$$\mathbf{E}\left[X \right] = \sum_{i=1}^{m} \mathbf{E}\left[X_i \right] = m \cdot \mathbf{Pr}\left[\mathcal{E}_i \right] = m \cdot \left(1 - \frac{1}{8} \right) = \frac{7m}{8}.$$

(b) Let k_i be the number of literals in the i-th clause. Again, let X_i be the indicator of the event $\mathcal{E}_i$ that the i-th clause is satisfied. Then,

$$\mathbf{Pr}\left[\mathcal{E}_i \right] = 1 - 2^{-k_i}.$$

Hence, the total number X of clauses satisfied is in expectation,

$$\mathbf{E}\left[X \right] = m - \sum_{i=1}^{m} 2^{-k_i}.$$

Exercise 21 [Pattern Matching] Consider a fixed pattern $P = (p_1, \ldots, p_k)$ with characters from an alphabet Σ. We want to find the expected number of times that this pattern occurs in a string X of length n whose characters are chosen uniformly at random from Σ. (For instance, the pattern "`abba`" appears three times in "`abaabbabbacabbacc`")

(*Answer*) If $k > n$, then the answer is 0. Otherwise, there are $n - k + 1$ possible starting points for pattern P in X. For each of these positions we define Z_i to be the indicator of the event $\mathcal{E}_i$ that P matches the string $X_{i:i+|P|-1}$.

Then, each of the characters of X has a probability $1/|\Sigma|$ to match the corresponding of P. Hence,

$$\mathbf{Pr}\left[\,\mathcal{E}_i\,\right] = \frac{1}{|\Sigma|^k}.$$

Hence, the expected number of occurrences of P in X is

$$\mathbf{E}\left[\,Z\,\right] = \sum_{i=1}^{n-k+1} \mathbf{E}\left[\,Z_i\,\right] = (n-k+1) \cdot \frac{1}{|\Sigma|^k}.$$

Question: What would change if the pattern P was not fixed?

Exercise 22 [LCS Expectation] Consider two strings X and Y of length n whose characters are generated uniformly at random from the alphabet Σ. Show that the expected length of the LCS of these two strings is $\Theta(n)$.

(*Answer*) An upper bound of n is trivial. For the lower bound, we can show that with high probability $c_1 n$ of the first characters of X appear as a subsequence in Y.

We do this by letting W_i be the time until we see character X_i in string Y given that we have seen a subsequence of $X_1, \ldots, X_{i-1}$ in Y. Then, the W_i's are independent Geometric random variables with success probability $p = \frac{1}{|\Sigma|}$.

Exercise 23 [LIS Expectation Lower bound] In this exercise, you will prove that the length of the longest increasing subsequence in a random permutation of $[n] := \{1, \ldots, n\}$ is $\Omega(\sqrt{n})$.

(a) Consider any sequence X of length n. Further, let ℓ_i be the length of the LIS ending at character i and similarly u_i be the length of the longest decreasing subsequence ending at i.

   i. Show that for any $i < j$ we must have that $(\ell_i, u_i) \neq (\ell_j, u_j)$.

   ii. By considering the pairs with values (ℓ_i, u_i) such that $0 \leq \ell_i, u_i \leq \sqrt{n}/2$, argue that there must be an increasing or decreasing sequence with length at least $\sqrt{n}/2$.

(b) Deduce that in a random permutation, the length L of the longest common subsequence satisfies

$$\mathbf{E}\left[\,L\,\right] \geq \frac{\sqrt{n}}{4}.$$

(*Answer*)

(a) i. Consider $i < j$ with $\ell_i = \ell_j$ and $u_i = u_j$. Then we consider two cases:

- **Case $[X_i < X_j]$:** Then, we can extend the longest increasing subsequence at i by appending X_j (since $X_i < X_j$) and so $\ell_j \geq \ell_i + 1$. (Contradiction)
- **Case $[X_i > X_j]$:** Then, we can extend the longest decreasing subsequence at i by appending X_j (since $X_j < X_i$) and so $u_j \geq u_i + 1$. (Contradiction)

Hence, we cannot have both $\ell_i = \ell_j$ and $u_i = u_j$.

ii. by i., there are in total n distinct pairs. Also, there are $(\sqrt{n}/2)^2 = n/4$ pairs with $0 \leq \ell_i, u_i \leq \sqrt{n}/2$.

(b) By the analysis in (a), every sequence X has a LIS or LDS of length at least $\sqrt{n}/2$. Since LIS and LDS is symmetric this implies that at least half of the permutations have a LIS of length at least $\sqrt{n}/2$. Hence,

$$\mathbf{E}\left[\,L\,\right] \geq \frac{\sqrt{n}}{4}.$$

1.4 Useful identities

Exercise 24 [Expectation as sum of probabilities]

(a) Consider a discrete random variable $X \geq 0$. Prove that

$$\mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{Pr}[X \geq i].$$

(b) Use the above formula to derive the expectation of the geometric random variable $X \sim \mathsf{Geom}(p)$.

2 Probabilistic Method

Exercise 25 Show that every graph $G = (V, E)$ contains a bipartite subgraph with at least $|E|/2$ edges.

3 Various Algorithms

Exercise 26 [Matrix Multiplication Verification] Given three $n \times n$ matrices A, B, C, we want to check if $A \times B = C$. The best known matrix multiplication algorithm takes time $\Omega(n^{2.37})$, so we would like to do the verification faster than that. For simplicity assume that the matrices are binary (i.e., their entries are just $0, 1$) and the operations are those modulo 2. Consider the algorithm that randomly samples a vector $v \in \{0, 1\}^n$ and outputs 1 iff $(A \times B)v = Cv$.

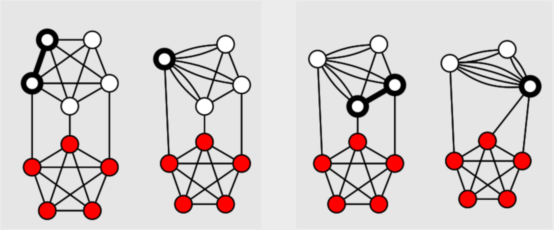
(a) Argue that the check $(A \times B)v = Cv$ can be performed more efficiently than matrix multiplication.

(b) Show that the algorithm succeeds with probability $\geq 1/2$.

(c) How would you improve the success probability?

(d) How would you apply the algorithm for other inputs?

Exercise 27 [Randomised Min-Cut] In this exercise, you will analyse a simple randomised algorithm for finding the minimum cut in an undirected graph. The algorithm proceeds by sampling a random edge (u, v) in the graph and *contracting* its endpoints. The figures below give two examples of contractions:



The algorithm makes $n - 2$ contractions until there are two vertices remaining. The number of edges between the two vertices will the answer for the minimum cut.

(a) Show that if the minimum cut in the graph has size M, then each vertex has degree at least M.

(b) Fix a minimum cut C. Let E_i be the event that in the i-th step the algorithm did not contract an edge of C. Show that

$$\mathbf{Pr}[E_1] \geq 1 - \frac{2}{n} \quad \text{and} \quad \mathbf{Pr}\left[E_i \;\middle|\; \bigcap_{j=1}^{i-1} E_j\right] \geq 1 - \frac{2}{n - i + 1}.$$

(c) Prove that

$$\mathbf{Pr}\left[\bigcap_{j=1}^{n-2} E_j\right] \geq 1 - \frac{2}{n \cdot (n-1)}.$$

(d) Using amplification, design an algorithm with success probability at least $1 - n^{-1}$ of finding a minimum cut. What is the running time of your algorithm?

(*Answer*)

(a) If there was a vertex $v \in V$ with degree less than M, there there would be a cut $S = \{v\}$ (which just disconnects v) with a size smaller than the min-cut (which is a contradiction). Hence, all vertices have degree at least M.

(b) Since each vertex has degree at least M, there must be a total of at least $nM/2$ edges. Hence, assuming that we have not contracted any of the min-cut edges, i.e., $\bigcap_{j=1}^{i-1} E_j$ holds, then there are $n - i$ vertices in the graph so

$$\mathbf{Pr}\left[E_i \;\middle|\; \bigcap_{j=1}^{i-1} E_j \right] \geq 1 - \frac{M}{(n-i+1)M/2} = 1 - \frac{2}{n-i+1}.$$

(c) Using the chain rule for probability events and the result from (b) we have that

$$\begin{aligned}\mathbf{Pr}\left[E_i \;\middle|\; \bigcap_{j=1}^{i-1} E_j \right] &= \mathbf{Pr}\left[\bigcap_{j=1}^{n-2} E_j \;\middle|\; \bigcap_{j=1}^{n-3} E_j \right] \cdot \mathbf{Pr}\left[\bigcap_{j=1}^{n-3} E_j \;\middle|\; \bigcap_{j=1}^{n-4} E_j \right] \cdot \ldots \cdot \mathbf{Pr}\left[E_2 \mid E_1 \right] \cdot \mathbf{Pr}\left[E_1 \right] \\ &\geq \prod_{j=1}^{n-2} \frac{n-j-1}{n-j+1} \\ &= \frac{1}{3} \cdot \frac{2}{4} \cdot \ldots \cdot \frac{n-4}{n-2} \cdot \frac{n-3}{n-1} \cdot \frac{n-2}{n} \\ &= \frac{1}{n \cdot (n-1)},\end{aligned}$$

by cancelling out terms from the numerators and denominators.

(d) By repeating the algorithm $t := \frac{n \cdot (n-1)}{2} \cdot \log n$ times, we get probability at least $1 - n^{-1}$ (see Exercise 4).

Exercise 28 [All min-cuts]

(a) Using Exercise 27 (c), argue that any graph can have at most $\binom{n}{2}$ min cuts.

(b) (optional) Show that there is a graph with this number of cuts.

(c) Modify the algorithm from Exercise 27 to return all minimum cuts for the given graph.

(*Answer*)

(a) In Exercise 27 (c), we proved that any cut C can be found using this randomised algorithm with success probability at least $\frac{1}{n \cdot (n-1)}$.

Now assume there are k minimum cuts $C_1, \ldots, C_k$, then the events E_i (for $1 \leq i \leq k$) that the algorithm finds cut C_i are disjoint, as after each run the algorithm produces a single cut. Hence,

$$\mathbf{Pr}\left[C_1 \cup \ldots \cup C_k \right] = \sum_{i=1}^{k} \mathbf{Pr}\left[C_i \right] \geq \frac{2k}{n \cdot (n-1)}.$$

Since any probability is upper bounded by 1, we have that

$$k \leq \frac{n \cdot (n-1)}{2} = \binom{n}{2},$$

which concludes the claim.

(b) The cycle C_n with n vertices is such a graph. To disconnect the cycle, we need to remove any two of the n edges. There are $\binom{n}{2}$ ways of selecting these two edges, which matches the upper bound.

(c) Consider the following algorithm A', letting A be a single iteration of the algorithm in Exercise 27. Run A for at least $t = 6n \cdot (n-1) \cdot \log n$ steps and return all the cuts that achieved the minimum value.

By the analysis in 27 (c), it follows that the probability of finding a minimum cut C_i in any repetition is at least $1/(n \cdot (n-1))$. Hence, the probability of the event E_i of finding C_i in at least one repetition is

$$\mathbf{Pr}\left[\, E_i \,\right] \geq 1 - n^{-3}.$$

By (a), there are at most $M \leq n \cdot (n-1)/2$ cuts. So, the probability that we don't find any of them is upper bounded by

$$\mathbf{Pr}\left[\bigcap_{i=1}^{M} \neg E_i\right] \leq M \cdot n^{-3} \leq n^{-1}.$$

Hence, the algorithm A' succeeds with probability at least $1 - n^{-1}$.

Exercise 29 [Karger-Stein algorithm] (++) Consider the following improvement to the algorithm in Exercise 27. For convenience we define the function $\texttt{Contract}(G, t)$ to perform t random edge contractions on graph G.

```
function FasterMinCut($G = (V, E)$)
    if $|V| \leq 6$ then
        return BruteForceMinCut($G$)
    else
        $t \leftarrow \lceil 1 + |V|/\sqrt{2} \rceil$
        $G_1 \leftarrow$ Contract($G, t$)
        $G_2 \leftarrow$ Contract($G, t$)
    end if
    return min(FasterMinCut($G_1$), FasterMinCut($G_2$))
end function
```

(a) Show that the running time of $\texttt{FasterMinCut}(G)$ for a graph with n vertices satisfies:

$$T(n) = 2T\left(\lceil 1 + n/\sqrt{2} \rceil\right) + \mathcal{O}(n^2).$$

(b) Show that $T(n) = \mathcal{O}(n^2 \log n)$.

(c) Show that a lower bound $P(n)$ on the probability of success for a graph with n vertices, satisfies the following recurrence

$$P(n) = 1 - \left(1 - \frac{1}{2} P\left(\lceil 1 + n/\sqrt{2} \rceil\right)\right).$$

(d) (++) Show that $P(n) = \Omega(\log n)$.

(e) Obtain an algorithm for min-cut with success probability at least $1 - n^{-1}$ and running time $\mathcal{O}(n \log^3 n)$.

Exercise 30 [Random subsets] You are given a finite set $S = \{s_1, \ldots, s_n\}$.

(a) Give an algorithm to generate a random subset of S.

(b) Let X and Y be two random subsets of S.

  i. What is the probability that $X \subseteq Y$?

  ii. What is the probability that $X \cup Y = S$?

  iii. What is the expected size of $X \cup Y$?

  iv. What is the expected size of $X \cap Y$?

(*Answer*)

(a) For each element s_i of S, independently toss an unbiased coin and if the outcome is heads, then add s_i to the set, otherwise don't add. The probability that any fixed subset $X \subset S$ is sampled is $2^{-1} \cdot \ldots \cdot 2^{-1} = 2^{-n}$, i.e., uniform among the 2^n possible subsets.

(b) i. Fix a set $x \subseteq S$ with k elements. Then, Y will be a superset of x if for each of these k elements the tosses were heads (the rest of the elements do not matter). Therefore,

$$\mathbf{Pr}\left[\, \{X = x\} \cap \{X \subset Y\} \,\right] = \frac{1}{2^k} \cdot \frac{1}{2^{n-k}} \cdot \frac{1}{2^k} = \frac{1}{4^k} \cdot \frac{1}{2^{n-k}}.$$

Hence, using the binomial theorem

$$\mathbf{Pr}\left[\, X \subseteq Y \,\right] = \sum_{k=1}^{n} \binom{n}{k} \cdot \frac{1}{4^k} \cdot \frac{1}{2^{n-k}} = \left(\frac{1}{4} + \frac{1}{2}\right)^n = 0.75^n.$$

ii. Similarly to the previous question, fix a set $x \subseteq S$ with k elements. The set Y should have heads in all remaining 2^{n-k} values. Therefore,

$$\mathbf{Pr}\left[\, \{X = x\} \cap \{X \cup Y = S\} \,\right] = \frac{1}{2^k} \cdot \frac{1}{2^{n-k}} \cdot \frac{1}{2^{n-k}} = \frac{1}{2^k} \cdot \frac{1}{4^{n-k}}.$$

Hence, using again the binomial theorem gives

$$\mathbf{Pr}\left[\, X \subseteq Y \,\right] = \sum_{k=1}^{n} \binom{n}{k} \cdot \frac{1}{2^k} \cdot \frac{1}{4^{n-k}} = \left(\frac{1}{2} + \frac{1}{4}\right)^n = 0.75^n.$$

iii. Let Z_i be the indicator random variable for whether $s_i \in X \cap Y$. Then,

$$\mathbf{E}\left[\, |X \cap Y| \,\right] = \sum_{i=1}^{n} \mathbf{E}\left[\, Z_i \,\right] = n \cdot \frac{1}{4} = \frac{n}{4},$$

since $\mathbf{Pr}\left[\, Z_i \,\right] = 1/4$, the probability that both coin tosses are in heads.

iv. Similarly, to the previous question, let Let Z_i be the indicator random variable for whether $s_i \in X \cap Y$. Then,

$$\mathbf{E}\left[\, |X \cup Y| \,\right] = \sum_{i=1}^{n} \mathbf{E}\left[\, Z_i \,\right] = n \cdot \frac{1}{4} = \frac{3}{4}n,$$

since $\mathbf{Pr}\left[\, Z_i \,\right] = 3/4$, the probability that at least one of the coin tosses are in heads.

Exercise 31 [Reservoir sampling] We are given a stream of N values, where N is large and unknown. The values will be presented one by one and we can only store one value at each step (perhaps because we have limited memory).

We want to sample from this stream one value uniformly at random.

(a) How would you sample the value if N was known?

(b) Consider an algorithm that stores the first value in memory and then for the k-th value a_k with probability $1/k$ it replaces the value in memory with a_k and otherwise it keeps it unchanged.

Show that this algorithm produces a uniform sample from the stream.

(c) Generalise the previous algorithm so that it samples M values uniformly at random from the stream. At each point in time your algorithm can keep at most M values in memory.

(*Answer*)

(a) We could uniformly sample an index i from 1 to N and keep a counter of the number of samples encountered so far. When the counter reaches i, then we return this as our sample.

(b) Let $\mathcal{E}_k$ be the event that we report value a_k as the sample. For this to happen, (i) we must place a_k in memory (happens with probability $1/k$) and (ii) none of the later values should replace it. This happens with probability

$$\mathbf{Pr}\left[\, \mathcal{E}_k \,\right] = \frac{1}{k} \cdot \left(1 - \frac{1}{k+1}\right) \cdot \left(1 - \frac{1}{k+2}\right) \cdot \ldots \cdot \left(1 - \frac{1}{N}\right) = \frac{1}{k} \cdot \frac{k}{k+1} \cdot \frac{k+1}{k+2} \cdot \ldots \cdot \frac{N-1}{N} = \frac{k}{N}.$$

(c) We modify the algorithm so that value a_k is added to the store with probability M/k and then a value out of the M available is replaced uniformly. Therefore, for $\mathcal{E}_k$ being the event that we report a_k in the end, we have that

$$\mathbf{Pr}\left[\, \mathcal{E}_k \,\right] = \frac{M}{k} \cdot \left(1 - \frac{M}{k+1} \cdot \frac{1}{M}\right) \cdot \left(1 - \frac{M}{k+2} \cdot \frac{1}{M}\right) \cdot \ldots \cdot \left(1 - \frac{M}{N} \cdot \frac{1}{M}\right) = \frac{M}{N},$$

which shows that the algorithms computes an unbiased estimate.

Exercise 32 [Algorithm L] Read about the Algorithm L for reservoir sampling and answer the following questions:

(a) Explain how this algorithm works (without having knowledge of the number of elements in the stream).

(b) In What sense is it better than the algorithm described in Exercise 31?

Exercise 33 [Algorithm A-Res] Read about Algorithm A-Res and answer the following questions:

(a) Describe the reservoir problem with weights.

(b) Explain how this algorithm works and prove its correctness.

Exercise 34 [Searching a Random Hash Table] In the **Lecture 2 slide 12**, it was shown that if we throw n balls into n bins, then the maximum load of any bin is at most $4 \cdot \log n / \log\log n$ with probability at least $1 - n^{-1}$.

In this exercise, we are interested in upper bounding the number of operations required for inserting n (random) elements in a hash table. The additional work is that when we insert an element v to a bin i with X_i elements, then we need to go over all of its elements to check if v is present. Therefore, this requires $\mathcal{O}(X_i)$ time.

(a) Using the analysis in the lecture, argue that the total time to insert n random keys to a hash table is $\mathcal{O}(n \cdot (\log n / \log\log n)^2)$ with high probability.

(b) Prove that for a binomial r.v. $N \sim \textsc{Bin}(n, p)$,

$$\mathbf{E}\left[\, N^2 \,\right] = n \cdot p + n \cdot (n-1) \cdot p^2.$$

(c) Using (b), show that the amortised insert time is $\mathcal{O}(n)$ in expectation.

(*Answer*)

(a) By the analysis in the lecture notes we have that $\max_{i \in [n]} X_i \leq 4 \cdot (\log n / \log\log n)^2$ for some constant $c > 0$ with probability at least $1 - n^{-1}$. When $X_i = x$, then the total insert time for these elements is at most:

$$1 + 2 + \ldots + x = \frac{1}{2} x \cdot (x+1) \leq x^2.$$

Hence, on aggregate we get that with probability at least $1 - n^{-1}$, the total work for the insertion of n elements is at most

$$\frac{1}{2} \cdot \sum_{i=1}^{n} X_i^2 \leq 8n \cdot \left(\frac{\log n}{\log\log n} \right)^2.$$

(b) By writing $N = \sum_{j=1}^{n} Z_j$, where $Z_j \sim \textsc{Ber}(p)$, we have that

$$\begin{aligned}
\mathbf{E}\left[\, N^2 \,\right] &= \mathbf{E}\left[\left(\sum_{j=1}^{n} Z_j \right)^2 \right] \\
&\leq \sum_{j=1}^{n} \mathbf{E}\left[\, Z_j^2 \,\right] + \sum_{j=1}^{n} \sum_{k=1, k \neq j}^{n} \mathbf{E}\left[\, Z_j Z_k \,\right] \\
&\leq \sum_{j=1}^{n} \mathbf{E}\left[\, Z_j \,\right] + \sum_{j=1}^{n} \sum_{k=1, k \neq j}^{n} \mathbf{E}\left[\, Z_j \,\right] \mathbf{E}\left[\, Z_k \,\right] \\
&\leq n \cdot p + n \cdot (n-1) \cdot p^2.
\end{aligned}$$

(c) Using (b) for $p = 1/n$, we get that

$$\mathbf{E}\left[\, X_i^2 \,\right] = 1 + 1 - \frac{1}{n} \leq 2.$$

Hence, in expectation the total amount of work is

$$\mathbf{E}\left[\sum_{i=1}^{n} X_i^2\right] \leq 2n.$$

4 Concentration inequalities

Exercise 35 [Chernoff Bound for Balls-into-Bins] Use the following Chernoff bound

$$\mathbf{Pr}\left[\, X \geq (1+\delta) \cdot \mathbf{E}\left[\, X \,\right] \,\right] \leq \exp\left(-\frac{\delta^2 \mathbf{E}\left[\, X \,\right]}{3}\right),$$

with an appropriate choice of δ to show that when allocating m balls into n bins uniformly at random, with $m \geq n \log n$, then the the maximum load of any bin is at most $\frac{m}{n} + \mathcal{O}(\sqrt{\frac{m}{n}} \cdot \log n)$ with probability at least $1 - n^{-2}$.

(*Answer*) Let X_i be the load of bin $i \in [n]$ after the m allocations. Then $X_i = \sum_{j=1}^{m} Z_j$ where $Z_j \in \{0, 1\}$ indicates whether the j-th ball was allocated to bin i. Since the balls are allocated uniformly we have that $\mathbf{Pr}\left[\, Z_j \,\right] = \frac{1}{n}$ and also that $\mathbf{E}\left[\, X_i \,\right] = \frac{m}{n}$.

Now, how should we pick δ? We want to pick δ such that $(1+\delta) \cdot \frac{m}{n}$ is $\frac{m}{n} + \mathcal{O}(\sqrt{\frac{m}{n}} \cdot \log n)$. Therefore, we pick

$$\delta := C \cdot \sqrt{\frac{m}{n} \cdot \log n},$$

where $C > 0$ is a constant that we will choose later (to get the probabilities to work correctly). Therefore, applying the Chernoff bound we have that

$$\mathbf{Pr}\left[X \geq \frac{m}{n} + C \cdot \sqrt{\frac{m}{n} \cdot \log n} \right] \leq \exp\left(-\frac{C^2 \log n}{3}\right),$$

By setting $C := 3$, we have that

$$\mathbf{Pr}\left[X_j \geq \frac{m}{n} + C \cdot \sqrt{\frac{m}{n} \cdot \log n} \right] \leq n^{-3}.$$

Now, it remains to prove that this bound holds for all bins. For this we take the union bound over the n bins,

$$\mathbf{Pr}\left[\max_{j \in [n]} X_j \geq \frac{m}{n} + 3 \cdot \sqrt{\frac{m}{n} \cdot \log n} \right] \leq n^{-2}.$$

Exercise 36 [Number of inversions Concentration] In Exercise 15, you proved that for a random permutation of elements $[n] := \{1, \ldots, n\}$, the number of inversions X is $\frac{1}{4} \cdot n(n-1)$ in expectation. In this exercise, you will apply McDiarmid's inequality to obtain a concentration bound.

(a) Argue that one can generate a random permutation of $[n]$ by sampling n independent values from $\mathcal{U}[0, 1]$ and then replacing each with their rank (e.g., $(0.1, 0.41, 0.23, 0.21)$ would correspond to $(1, 4, 3, 2)$),

(b) Let f be the function that counts the number of inversions in a given permutation. Prove that f is Lipschitz with parameter n.

(c) Use McDiarmid's inequality to deduce that

$$\mathbf{Pr}\left[\left| X - \frac{1}{4} \cdot n(n-1) \right| < \sqrt{n^3 \log n} \right] \geq 1 - 2n^{-2}.$$

(*Answer*)

(a) We start by noting that the probability two samples form $\mathcal{U}[0, 1]$ to be equal is 0. Then, because the n random variables are completely symmetric each one is equally likely to be at any fixed rank i. So the rank of these elements creates a random permutation.

(b) Let $f : [0,1]^n \to \mathbb{N}$ be the function that counts the number of inversions. Then, for any $u_1, \ldots, u_n$, any position i and any $u_i' \in [0,1]$, we have that

$$|f(u_1, \ldots, u_{i-1}, u_i, u_{i+1}, \ldots, u_n) - f(u_1, \ldots, u_{i-1}, u_i', u_{i+1}, \ldots, u_n)| \leq n,$$

since changing one element can increase the number of inversions by at most n.

(c) Since the $u_1, \ldots, u_n$ are independent and f is 1-Lipschitz, we have that

$$\mathbf{Pr}\left[\,|f - \mathbf{E}[\,f\,]| > t\,\right] \leq 2 \cdot \exp\left(-\frac{2t^2}{n \cdot n^2}\right).$$

By choosing $t := \sqrt{n^3 \log n}$, we get the conclusion

$$\mathbf{Pr}\left[\,|f - \mathbf{E}[\,f\,]| > \sqrt{n^3 \log n}\,\right] \leq 2 \cdot \exp\left(-\frac{2n^3 \log n}{n \cdot n^2}\right) = 2n^{-2}.$$

Exercise 37 [Number of Records Concentration] In Exercise 16, you proved that the expected number of records in a random permutation of elements in $[n]$ is $\log n + \Theta(1)$ in expectation. In this exercise, you will prove concentration.

(a) Let $\mathcal{E}_i$ be the event that element i is smaller than all previous elements. Then for any $i \neq j$ the events $\mathcal{E}_i$ and $\mathcal{E}_j$ are independent.

(b)

(*Answer*)

(a) We will generate the random permutation by assigning ranks to elements one by one starting with $i = 1$, then $i = 2$ and so on. The i-th element has $(i-1)+1$ possible ranks

(b) The number of records is a function of the

Exercise 38 [Local Maxima Concentration] In Exercise 17, you proved that the expected number of local maxima in a random permutation of $[n] := \{1, \ldots, n\}$ is $(n+1)/3$.
Prove that the number of local maxima is concentrated around this mean.

(*Answer*) We generate the random permutation by first generating n independent uniform random variables $U_1, \ldots, U_n \sim \mathcal{U}[0,1]$ and then reporting the rank of the values as the permutation.
Then the number of local maximuma is a function $f(U_1, \ldots, U_N)$. By changing any one value, we can change the number of local maxima by at most 2 (*why not* 3*?*).
Hence, applying McDiarmid's inequality, we have that

$$\mathbf{Pr}\left[\,|f - \mathbf{E}[\,f\,]| \geq t\,\right] \leq 2 \cdot \exp\left(-\frac{2t^2}{n \cdot 2^2}\right).$$

By setting $t := 2\sqrt{\log n}$, we get that

$$\mathbf{Pr}\left[\,|f - \mathbf{E}[\,f\,]| < 2\sqrt{\log n}\,\right] \geq 1 - n^{-1}.$$

Exercise 39 [Longest Increasing Subsequence Concentration (I)] In this exercise, you will prove that the length of the longest increasing subsequence is $\mathcal{O}(\sqrt{n})$ with high probability.

(a) Let $X_{n,k}$ be the number of increasing subsquences of length k. Show that

$$\mathbf{E}[\,X_{n,k}\,] = \frac{1}{k!} \cdot \binom{n}{k}.$$

(b) Using Exercise 7, show that

$$\mathbf{E}[\,X_{n,k}\,] \leq \frac{n^k}{(k/e)^{2k}}.$$

(c) By arguing that

$$\mathbf{Pr}\left[\, L \geq k \,\right] \leq \mathbf{E}\left[\, X_{n,k} \,\right],$$

and using Markov's inequality, show that for any constant $C > 2$, we have that

$$\mathbf{Pr}\left[\, L \geq Ce\sqrt{n} \,\right] \leq \left(\frac{1}{C}\right)^{2Ce\sqrt{n}}.$$

(d) Deduce that $\mathbf{E}\left[\, L \,\right] = \mathcal{O}(\sqrt{n})$.

Exercise 40 [Longest Increasing Subsequence Concentration (II)] In Exercise 39, you proved that the length of the longest subsequence in a random permutation of $[n]$ is $\Theta(\sqrt{n})$ in expectation. As an alternative, use McDiarmid's inequality to prove that it is concentrated around the expectation.

(*Answer*) We will generate the random permutation by sampling n uniform random variables $u_1, \ldots, u_n \sim \mathcal{U}[0,1]$ and then generate a permutation by looking at their ranks.
Let $f : [0,1]^n \to \mathbb{N}$ be the function that takes $u_1, \ldots, u_n$ and returns the length of the longest increasing subsequence of the corresponding permutation. We will now show that this function is 1-Lipschitz. Consider the change of any of the coordinates i from u_i to u_i'. Then, for any increasing subsequence $i_1 < \ldots < i_k$, its length can either increase by 1 (if i was not in the subsequence) or it decreases by 1 (if i was in the subsequence and now violates the central property). Therefore,

$$|f(u_1, \ldots, u_{i-1}, u_i, u_{i+1}, \ldots u_n) - f(u_1, \ldots, u_{i-1}, u_i', u_{i+1}, \ldots, u_n)| \leq 1.$$

By applying McDiarmid's inequality, we have that

$$\mathbf{Pr}\left[\, |f - \mathbf{E}\left[\, f \,\right]| \geq t \,\right] \leq 2 \cdot \exp\left(-\frac{2t^2}{n \cdot 1^2}\right)$$

By choosing $t = \Theta(\sqrt{n})$ or $t = \Theta(\sqrt{n \log n})$, we get concentration.

Exercise 41 [Pattern Matching Concentration] In Exercise 21, we found the expectation for the number of occurrences N of pattern P (of length k) in a random string X (of length n).

(a) Prove that N is concentrated around its mean.

(b) For which values of N and k does your bound make sense?

(*Answer*)

(a) Let $f : \Sigma^n \to \mathbb{N}$ be the function that takes as input the n random characters of X and outputs the number of occurrences of P in X. Any character appears in at most k substrings of length k. Hence, by changing one character, we can change the number of occurrences by at most k. Therefore, f satisfies the k-Lipschitz property

$$\mathbf{Pr}\left[\, |f - \mathbf{E}\left[\, f \,\right]| \geq t \,\right] \leq 2 \cdot \exp\left(-\frac{2t^2}{n \cdot k^2}\right).$$

(b) Note that if $k > 3\log_{|\Sigma|} n$, then the number of occurrences is < 1 in expectation. For values less than that, we can choose $t = k\sqrt{n \log n}$ to get a concentration with high probability (and the value of t will be small in comparison to the expectation).

Exercise 42 [Searching a Random Hash Table Concentration] In Exercise 34, you proved that the insertion time of n random elements in a hash table takes amortised $\mathcal{O}(n)$ time in expectation.
Let X_i be the loads of the bins after the n elements have been allocated, $Y_i = X_i^2$ and $Y_i' = \min(Y_i, \log^2 n)$. Further let $Y = \sum_{i=1}^n Y_i$ and $Y' = \sum_{i=1}^n Y_i'$.

(a) Using Exercise 34, argue that $\mathbf{E}\left[\, Y' \,\right] = \mathcal{O}(n)$.

(b) Argue that Y' is $(5\log n)$-Lipschitz with respect to the n elements being inserted.

(c) Prove that Y' is concentrated around its mean.

(d) Deduce that Y is concentrated.

(*Answer*)

(a) Note that $Y_i' \leq Y_i$ and so $Y' \leq Y$. Therefore, by Exercise 34, we have that $\mathbf{E}[Y'] \leq \mathbf{E}[Y] \leq 2n$.

(b) By moving one ball from bin i to bin j

$$\begin{aligned}|\Delta Y'| &\leq |(\min(X_i - 1, \log n))^2 - Y_i'| + |(\min(X_j + 1, \log n))^2 - Y_j'| \\ &\leq 2 \cdot \left|(\log n + 1)^2 - (\log n)^2\right| \\ &\leq 2 \cdot (2 \log n + 1) \leq 5 \log n.\end{aligned}$$

(c) Since each ball is allocated independently we can apply McDiarmid's inequality to get

$$\mathbf{Pr}[\,|Y' - \mathbf{E}[Y']| \geq t\,] \leq 2 \cdot \exp\left(-\frac{2t^2}{n \cdot (5 \log n)^2}\right).$$

By choosing $t = n$, we (quite generously) get that

$$\mathbf{Pr}[\,Y' \leq 3n\,] \geq 1 - 2 \cdot \exp\left(-\frac{2n^2}{25n \log^2 n}\right) \geq 1 - n^{-3}.$$

(d) From the lectures, we know that

$$\mathbf{Pr}\left[\max_{i \in [n]} X_i \leq 4 \cdot \frac{\log n}{\log \log n}\right] \geq 1 - n^{-1}. \tag{1}$$

When $X_i \leq \log n$, we have that $Y_i' = Y_i$ and so

$$\mathbf{Pr}[\,Y = Y_i'\,] \geq \mathbf{Pr}\left[\max_{i \in [n]} X_i \leq 4 \cdot \frac{\log n}{\log \log n}\right] \geq 1 - n^{-1}.$$

By taking the union bound with 1, we have that

$$\mathbf{Pr}[\,Y \leq 3n\,] \geq 1 - n^{-1} - n^{-3} \geq 1 - 2n^{-1}.$$

Exercise 43 [LCS Concentration] In Exercise 22, you proved that the length n of the longest common subsequence between two random strings X and Y is $\Theta(n)$.
Prove that L is concentrated around its mean.

(*Answer*) Let $f(x_1, \ldots, x_n, y_1, \ldots, y_n)$ be the function that takes the two random strings and outputs the length of the longest common subsequence. Then by changing any of the characters the length of a common subsequence can either increase by 1, decrease by 1 or stay the same. Hence, f is 1-Lipschitz.
Since the characters of the two strings are chosen independently, we can apply McDiarmid's inequality, to get that

$$\mathbf{Pr}[\,|f - \mathbf{E}[f]\,| \geq t\,] \leq 2 \cdot \exp\left(-\frac{2t^2}{n}\right).$$

Hence, by choosing $t := \sqrt{n \log n}$ we have that

$$\mathbf{Pr}\left[|f - \Theta(n)| \geq \sqrt{n \log n}\right] \leq 2 \cdot n^{-2}.$$

Exercise 44 [Monte Carlo Estimation Concentration]

Exercise 45 [Chromatic Number Concentration] The *chromatic number* $\chi(G)$ of a graph is the minimum number of colours to colour the vertices of G such that no two adjacent vertices have the same colour.
Consider an undirected graph G where each edge is inserted independently with probability p. Show

that

$$\mathbf{Pr}\left[\,|\chi(G) - \mathbf{E}\left[\,\chi(G)\,\right]| \geq t\sqrt{n}\,\right] \leq 2 \cdot e^{-2t^2}.$$

(*Answer*) We define the random vectors $E_1, \ldots, E_n$, where E_i reveals the edges between vertex i and vertices $1, \ldots, i-1$. Note that E_i's are independent. Then the chromatic number of the graph is a function $f(E_1, \ldots, E_n)$.

Let's assume that we have changed the E_i's. Then, by assigning the colour of i to a new colour, the total number of colours increases by 1. Hence, f is 1-Lipschitz and so,

$$\mathbf{Pr}\left[\,|f - \mathbf{E}\left[\,f\,\right]| \geq t\sqrt{n}\,\right] \leq 2 \cdot \exp\left(-\frac{2t^2 n}{n \cdot 1^2}\right) \leq 2e^{-2t^2}.$$

Q: Why did we not define E_i to be the set of vertices adjacent to i? Because the E_i's would not be independent.

Q: Why did we not define E_{uv} for each edge (u, v)? Because there could be $\Theta(n^2)$ edges.

Exercise 46 [Isolated vertices] Consider a graph sampled uniformly at random from the set of graphs with n vertices and cn edges for $c > 0$ being constant. Let X be the number of vertices that are *isolated*, i.e., have zero degree.

(a) Compute the expected value of X.

(b) Prove concentration for X.

(*Answer*)

(a) Fix a vertex v. Let X_v be the indicator of the event $\mathcal{E}_v$ that vertex v is isolated. Then, this means that none of the edges got assigned to it. Letting $M = n \cdot (n-1)$, then

$$\mathbf{Pr}\left[\,\mathcal{E}_v\,\right] = \left(1 - \frac{n-1}{M}\right) \cdot \left(1 - \frac{n-1}{M-1}\right) \cdot \ldots \cdot \left(1 - \frac{n-1}{M-cn+1}\right).$$

(Note: This is roughly $e^{-(n-1)\cdot(\ln(M)-\ln(M-cn))}$)

5 Derandomisation

Exercise 47 [Derandomise Max-Cut] Derandomise the Max-cut algorithm that you saw in Lecture 1.

Exercise 48 [Derandomise 3-CNF] Derandomise the algorithm for the 3-CNF problem presented in Exercise 20.

6 Randomised data structures

Exercise 49 [Randomised Binary Search Trees] In Part IA, you saw that Binary Search Trees without some balancing mechanism (e.g., rotations) can lead to depths that are $\Omega(n)$ in size.

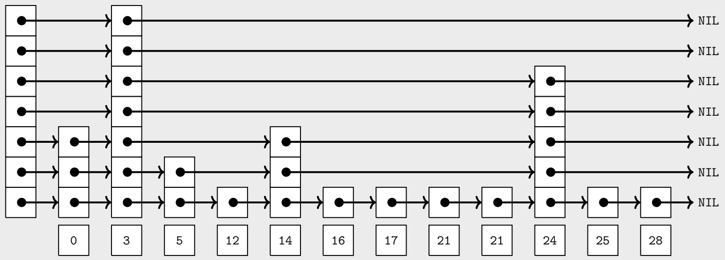
In this exercise you will analyse the expected time complexity for a BST where the input values are random.

(a) What is the expected number of comparisons when building a RBST?

(b) What is the expected number of comparisons when searching for a random element in the RBST?

Exercise 50 [SkipList] The SkipList is a data structure like lists, but it aims to support $\mathcal{O}(\log n)$ access time to every element with high probability and it can also serve as a BST.

The idea is that when inserting an element x we also add a random number of pointers, generated by sampling a Geom(1/2) (the 1/2 not being important). For instance, here is a randomly generated skip list, where element 14 has 3 pointers. To search for an element we start from the top left and follow the pointers as long as we are before our target element.



(a) Show that for all of the n items have at most $\mathcal{O}(\log n)$ levels with probability at least $1 - n^{-2}$.

(b) By arguing backwards from the target element to the root, show that the search algorithm needs $\mathcal{O}(\log n)$ heads in $\Theta(\log n)$ coin flips and prove that this happens with high probability.

(c) (optional +) Argue that in total this data structure uses $\mathcal{O}(n)$ memory.

7 Chernoff Bounds

Exercise 51 [For Geometric r.vs.] Consider n independent geometric random variables $X_1, \ldots, X_n$ with $X_i \sim \textsc{Geom}(p)$.

(a) Prove that for any $t > 0$,

$$\mathbf{E}\left[e^{tX_i} \right] = \frac{p}{e^{-t} - 1 + p}.$$

(b) Using the inequality $1 + x \leq e^x$, show that

$$\mathbf{E}\left[e^{tX_i} \right] \leq \frac{p}{p-t} = \left(1 - \frac{t}{p}\right)^{-1}.$$

(c) Compute the Chernoff bound for $X := \sum_{i=1}^{n} X_i$ and optimise the choice of t to get for any $\delta > 0$,

$$\mathbf{Pr}\left[X \geq (1+\delta)\mathbf{E}\left[X \right] \right] \leq e^{-n\cdot(\delta - \ln(1+\delta))}.$$

8 Puzzles

8.1 Generating Random Variables

Exercise 52 You are given a biased coin which produces heads with probability p and tails with probability $1 - p$, for some unknown $p \in (0, 1)$.

- Design an algorithm that uses samples from this biased coin and produces an unbiased binary value.

- How efficient is your algorithm, i.e. how many biased samples does it need in expectation to generate an unbiased one (as a function of p)?

Exercise 53

- How would you use a (unbiased) random binary variables to generate random integers in the set $\{0, 1, \ldots, n\}$

- What is the expected running time of your algorithm?

Further Reading 3 You can find more problems of this sort here.

Exercise 54 [Random permutation] Design an algorithm that samples a permutation of n elements uniformly at random. How efficient is your algorithm?